Bayesian Operational Modal Analysis with Asynchronous Data, Part I: Most Probable Value

Yi-Chen Zhu[1] and Siu-Kui Au[2]

Institute for Risk and Uncertainty and Centre for Engineering Dynamics

University of Liverpool, United Kingdom

Abstract

In vibration tests, multiple sensors are used to obtain detailed mode shape information about the tested structure. Time synchronisation among data channels is required in conventional modal identification approaches. Modal identification can be more flexibly conducted if this is not required. Motivated by the potential gain in feasibility and economy, this work proposes a Bayesian frequency domain method for modal identification using asynchronous 'output-only' ambient data, i.e. 'operational modal analysis'. It provides a rigorous means for identifying the global mode shape taking into account the quality of the measured data and their asynchronous nature. This paper (Part I) proposes an efficient algorithm for determining the most probable values of modal properties. The method is validated using synthetic and laboratory data. The companion paper (Part II) investigates identification uncertainty and challenges in applications to field vibration data.

Key Words: Ambient data, Asynchronous data, Bayesian methods, FFT, Operational modal analysis

1. Introduction

Modal identification aims at determining the modal characteristics of a structure, which primarily include natural frequencies, damping ratios and mode shapes [1,2]. These properties are demanded for industrial applications such as vibration control, damage detection and structural health monitoring [3–6]. Compared to traditional vibration tests based on free vibration (zero input) or forced vibration (known input), ambient vibration test

[1] Corresponding author. Harrison Hughes Building, Brownlow Hill, Liverpool, L69 3GH, UK. Email: sgyzhu7@liverpool.ac.uk.
[2] Email: siukuiau@liverpool.ac.uk

does not require artificial loading but assumes that it is broadband random. Based on random natural excitations such as wind, microtremor and cultural activities, ambient vibration test can be conducted during the daily use of structure. For its high economy and convenience in applications, it has attracted much interest in both theory development and field test applications over the past few decades [7–9].

Modal identification based on 'output-only' ambient vibration data is conventionally known as 'operational modal analysis' (OMA) [10,11]. Among other methods, frequency domain decomposition [12] provides a quick estimation based on sample power spectral density (PSD) [13]. Stochastic subspace identification [14–16] extracts modal parameters by least square estimation with a state-space model. Recently, transmissibility based OMA techniques have been proposed [17,18] with the premise of being robust to the characteristics of the excitation spectra in identifying mode shapes.

Bayesian approach views modal identification as a general inference problem based on available information. Methods [19,20] have been developed in different contexts, in the time domain [21], frequency domain based on sample PSD matrix [22–24] and Fast Fourier transform (FFT) of data [25,26]. Given information of data and modelling assumptions, the information about modal parameters is extracted using Bayes' theorem and encapsulated in the 'posterior' (i.e. given data) probability density function (PDF). For modal identification problem which is 'globally identifiable' [27], the PDF can be characterised by the most probable value (MPV, where it is peaked) and covariance matrix (reflecting identification uncertainty).

In order to obtain mode shape information, the vibrations at different locations of a structure are measured using multiple sensors. Conventional modal identification techniques require 'synchronous' data, where the digital data at different channels are sampled according to the same time scale. Synchronisation does not only mean that the data recorded at multiple channels should start at the same time but also at the same pace. Simply logging multiple channels of data on the same computer does not imply they are synchronised. Each data acquisition (DAQ) unit has its own clock for sampling, whose accuracy is affected by temperature, aging etc. [28]. When multiple data channels are recorded using independent DAQ units, they will not be perfectly synchronised.

Practically, synchronisation means that the sampling time difference between data channels is within a certain tolerance. In one conventional configuration, the analog signals of sensors

are sampled using a central synchronisation hardware. In full-scale tests, this requires long analogue cables, which inevitably introduce additional noise in the data. Alternative options are Network Time Protocol [29] through the internet or Global Positioning System [30] for outdoor applications. Wireless sensor networks have also been applied in vibration tests with synchronisation corrections [31–33].

Synchronisation comes with an overhead. The possibility of performing OMA with asynchronous data is worth exploring. Motivated by the above considerations, a Bayesian method is proposed in this work for modal identification using asynchronous output-only ambient data. Asynchronous data is generally a non-stationary process, which is difficult to model from first principles. A stationary model with imperfect coherence is proposed so that it is conducive to modal identification, while capturing the key asynchronous characteristics within suitable time scales. Based on this model, the likelihood function for Bayesian inference is derived and its mathematical structure is analysed. An algorithm is developed in this paper (Part I) for efficiently determining the MPV of modal parameters. The companion paper (Part II) focuses on efficient determination and investigation of identification uncertainty. Synthetic and laboratory data examples are presented to illustrate and verify the proposed method. An application to modal identification of a full-scale building is also presented.

This paper is organised as follow. Bayesian approach based on FFT of data in a general context is briefly reviewed in Section 2. An identification model for asynchronous data is presented in Section 3. Based on this model, the mathematical structure of the theoretical PSD matrix for asynchronous data is analysed in Section 4. To facilitate analysis and modal identification, simplifying assumptions are made and the resulting posterior PDF is derived in Section 5. An iterative algorithm for determining the MPV of modal parameters is developed in Section 6. The high signal-to-noise ratio asymptotic behaviour of the MPV is analysed in Section 7. The overall identification procedure is summarised in Section 8. The proposed algorithm for MPV is validated using synthetic and laboratory data in Section 9.

2. Bayesian Framework

The proposed modal identification method is based on Bayesian approach using the FFT of ambient data for probabilistic inference. The overall framework is briefly reviewed in this section. Input loading is unknown but assumed to be broadband random near the resonance

band of the modes of interest. Let $\{\hat{\ddot{\mathbf{x}}}_j \in R^n\}_{j=1}^N$ denote the measured ambient acceleration data at n degrees of freedom (DOF) of the subject structure; N is the number of samples per channel. It is assumed to consist of theoretical structural response $\ddot{\mathbf{x}}_j \in R^n$ under ambient excitation and prediction error $\boldsymbol{\varepsilon}_j \in R^n$:

$$\hat{\ddot{\mathbf{x}}}_j = \ddot{\mathbf{x}}_j + \boldsymbol{\varepsilon}_j \tag{1}$$

The prediction error accounts for the difference between the theoretical response and measured data, which may arise from measurement noise or modelling error. The 'scaled FFT', or FFT in short, of $\{\hat{\ddot{\mathbf{x}}}_j\}$ is defined as:

$$\mathcal{F}_k = \sqrt{\frac{2\Delta t}{N}} \sum_{j=1}^{N} \hat{\ddot{\mathbf{x}}}_j \exp\left[-2\pi \mathbf{i} \frac{(k-1)(j-1)}{N}\right] \tag{2}$$

where $\mathbf{i}^2 = -1$ and Δt is the sampling interval. Here, $\mathcal{F}_k$ corresponds to frequency $\mathrm{f}_k = (k-1)/N\Delta t\,(\mathrm{Hz})$ for k=1,..,$\mathrm{N_q}$, where $\mathrm{N_q}$ (integer part of N/2+1) is the index corresponding to the Nyquist frequency. Multiplying $\mathcal{F}_k$ by its conjugate transpose gives the sample PSD matrix. The scaling factor is defined such that the PSD is one-sided with respect to frequency in Hz. For modal identification, only the $\mathcal{F}_k$ within a selected frequency band dominated by the modes of interested is used.

Let $\boldsymbol{\theta}$ denote the set of modal parameters to be identified. Using Bayes' theorem with a uniform prior PDF, the posterior PDF of $\boldsymbol{\theta}$ given $\{\mathcal{F}_k\}$ is

$$p(\boldsymbol{\theta}|\{\mathcal{F}_k\}) \propto p(\{\mathcal{F}_k\}|\boldsymbol{\theta}) \tag{3}$$

where $p(\{\mathcal{F}_k\}|\boldsymbol{\theta})$ is called the 'likelihood function'. A uniform prior PDF is justified for modal identification because the typical data size is sufficiently large that the likelihood function is fast-varying compared to the pior PDF. Assuming that data is stochastic stationary, for long data duration and high sampling rate (i.e. large $N\Delta t$ and small Δt), it can be shown that $\{\mathcal{F}_k\}$ are asymptotically independent at different frequencies and jointly 'circularly complex Gaussian' [34]. The likelihood function is then given by:

$$p(\{\mathcal{F}_k\}|\boldsymbol{\theta}) = (\pi)^{-nN_f} \times \prod_k (\det \mathbf{E}_k)^{-1} \exp\left[-\sum_k \mathcal{F}_k^* \mathbf{E}_k^{-1} \mathcal{F}_k\right] \tag{4}$$

where '*' denotes conjugate transpose; $\mathbf{E}_k = E[\mathcal{F}_k \mathcal{F}_k^* | \boldsymbol{\theta}]$ is the theoretical PSD matrix and N_f is the number of FFT points in the selected frequency band. For analysis or computation, it is more convenient to work with the 'negative log-likelihood function' (NLLF):

$$L(\boldsymbol{\theta}) = \sum_k \ln \det \mathbf{E}_k + \sum_k \mathcal{F}_k^* \mathbf{E}_k^{-1} \mathcal{F}_k \tag{5}$$

such that

$$p(\boldsymbol{\theta} | \{\mathcal{F}_k\}) \propto \exp[-L(\boldsymbol{\theta})] \tag{6}$$

With sufficient data, modal identification problem is 'globally identifiable' [27]. The MPV of $\boldsymbol{\theta}$ can be determined by maximizing the posterior PDF, or equivalently minimizing the NLLF with respect to $\boldsymbol{\theta}$. The posterior PDF can then be approximated by a Gaussian PDF

$$p(\boldsymbol{\theta} | \{\mathcal{F}_k\}) = (2\pi)^{-n_\theta/2} (\det \hat{\mathbf{C}})^{-1/2} \exp\left[-\frac{1}{2} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \hat{\mathbf{C}}^{-1} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \right] \tag{7}$$

where n_θ is the number of parameters in $\boldsymbol{\theta}$; $\hat{\boldsymbol{\theta}}$ is the MPV and $\hat{\mathbf{C}}$ is the posterior covariance matrix, equal to the inverse of Hessian of NLLF at MPV.

3. Modelling Asynchronous Data

The theory in the last section applies to general ambient data, regardless of whether it is synchronous or not. The departure point for asynchronous data lies in the theoretical PSD matrix $\mathbf{E}_k$. Imperfect synchronisation in data can be due to the initial time shift (start time) between channels and random time drifts due to different sampling clocks. This paper focuses on the latter, which is the primary issue in real applications with multiple asynchronous DAQ units. The former (initial time shift) is assumed to be zero, as in reality it will be detected and compensated using conventional time delay estimation methods [35,36], after which the residual time delay is negligible compared to the random time drift. Random time drift accumulates over time, rendering asynchronous data non-stationary [37,38]. Generally, a non-stationary process is much more difficult to model or analyse than a stationary one. In this work, asynchronous data is modelled as a stationary process with imperfect coherence to facilitate analysis and modal identification, while capturing the key asynchronous characteristics within suitable time scales. Based on the model, the PSD matrix $\mathbf{E}_k$ of asynchronous data will be derived in this section.

For instructional purpose, the PSD matrix for synchronous data is first briefly reviewed. Consider a frequency band dominated by a single mode. Let $\boldsymbol{\varphi}$ denote the 'global mode shape' covering all measured DOFs. When the data at all measured DOFs are synchronised, the FFT of measured data within the band can be modelled as:

$$\mathcal{F}_k = \boldsymbol{\varphi}\ddot{\eta}_k + \boldsymbol{\varepsilon}_k \tag{8}$$

where $\ddot{\eta}_k$ denotes the FFT of modal acceleration $\ddot{\eta}(t)$; and $\boldsymbol{\varepsilon}_k$ is the scaled FFT of measurement noise. Assuming classically damped structure, $\ddot{\eta}(t)$ satisfies the modal equation of motion:

$$\ddot{\eta}(t) + 2\zeta\omega\dot{\eta}(t) + \omega^2\eta(t) = p(t) \tag{9}$$

where $\omega = 2\pi f$ (rad/s); f (Hz) and ζ are the natural frequency and damping ratio of the mode, respectively; $p(t)$ is the modal force. Assuming that the modal force and prediction error are independent and have respectively constant PSDs of S and S_e in the selected band, the PSD matrix is given by:

$$\mathbf{E}_k = E[\mathcal{F}_k\mathcal{F}_k^*] = SD_k\boldsymbol{\varphi}\boldsymbol{\varphi}^T + S_e\mathbf{I}_n \tag{10}$$

where

$$D_k = \left[\left(\beta_k^2 - 1\right)^2 + \left(2\zeta\beta_k\right)^2\right]^{-1} \qquad \beta_k = f/\mathrm{f}_k \tag{11}$$

is the dynamic amplification factor.

For asynchronous data, the modal responses contributing to different data channels need not follow exactly the same time variation. In this work, the test configurations on time synchronisation are assumed to be known, as is typically the case. Define a 'synchronous data group' as a set of data channels that sample the data synchronously (i.e. using the same clock). Let the whole measurement array comprise n_g groups. For a given mode shape $\boldsymbol{\varphi}$, let $\mathbf{u}_i \in R^{n_i}$ denote the part of $\boldsymbol{\varphi}$ measured by the ith group with n_i DOFs, i.e.

$$\boldsymbol{\varphi} = \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_{n_g} \end{bmatrix} \tag{12}$$

Let $\ddot{\eta}_{ki}$ and $\boldsymbol{\varepsilon}_{ki}$ be the FFT of the modal acceleration and prediction error associated with the i th group. The FFT of asynchronous data is modelled as:

$$\mathcal{F}_k = \begin{bmatrix} \mathbf{u}_1 \ddot{\eta}_{k1} \\ \vdots \\ \mathbf{u}_{n_g} \ddot{\eta}_{kn_g} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_{k1} \\ \vdots \\ \boldsymbol{\varepsilon}_{kn_g} \end{bmatrix} = \mathbf{U} diag\left\{\|\mathbf{u}_i\|\right\}_{i=1}^{n_g} \ddot{\boldsymbol{\eta}}_k + \begin{bmatrix} \boldsymbol{\varepsilon}_{k1} \\ \vdots \\ \boldsymbol{\varepsilon}_{kn_g} \end{bmatrix} \tag{13}$$

where $diag\left\{\|\mathbf{u}_i\|\right\}_{i=1}^{n_g}$ denotes a diagonal matrix with entries $\left\{\|\mathbf{u}_i\|\right\}_{i=1}^{n_g}$; $\ddot{\boldsymbol{\eta}}_k = [\ddot{\eta}_{k1},...,\ddot{\eta}_{kn_g}]^T$ and $\mathbf{U} \in R^{n \times n_g}$ is a block-diagonal matrix formed by the unit vectors $\left\{\mathbf{u}_i / \|\mathbf{u}_i\|\right\}_{i=1}^{n_g}$:

$$\mathbf{U} = \begin{bmatrix} \mathbf{u}_1 / \|\mathbf{u}_1\| & & \\ & \ddots & \\ & & \mathbf{u}_{n_g} / \|\mathbf{u}_{n_g}\| \end{bmatrix} \tag{14}$$

Assume that all data channels are set to start recording at the same time with the same number of sampling points. Each synchronous group uses its own clock for sampling. In this context, the modal contributions in different synchronous data groups are asynchronised realisations of the same modal response. It is therefore reasonable to assume that $\left\{\ddot{\eta}_{ki}\right\}_{i=1}^{n_g}$ are identically distributed. Then

$$E\left(\ddot{\eta}_{ki} \ddot{\eta}_{ki}^*\right) = SD_k \qquad i = 1,..., n_g \tag{15}$$

On the other hand, modelling the relationship of modal responses associated with different groups fundamentally is challenging because asynchronous data in the first place is a non-stationary process. Balancing model simplicity and utility, the asynchronous nature of data among different groups is empirically modelled through imperfect coherence:

$$\chi_{kij} = \frac{E\left(\ddot{\eta}_{ki} \ddot{\eta}_{kj}^*\right)}{\sqrt{E\left(\ddot{\eta}_{ki} \ddot{\eta}_{ki}^*\right) E\left(\ddot{\eta}_{kj} \ddot{\eta}_{kj}^*\right)}} \tag{16}$$

where $\chi_{kij} \in C$ ($\left|\chi_{kij}\right| \le 1$) is the coherence between the i th and j th group at frequency f_k.

Based on the above, the theoretical PSD matrix for asynchronous data is given by

$$\mathbf{E}_k = E[\mathcal{F}_k \mathcal{F}_k^*] = SD_k \mathbf{U} \mathbf{C}_k \mathbf{U}^T + \begin{bmatrix} S_{e1}\mathbf{I}_{n_1} & & \\ & \ddots & \\ & & S_{en_g}\mathbf{I}_{n_g} \end{bmatrix} \tag{17}$$

where

$$\mathbf{C}_k = diag\left\{\|\mathbf{u}_i\|\right\}_{i=1}^{n_g} \boldsymbol{\chi}_k diag\left\{\|\mathbf{u}_i\|\right\}_{i=1}^{n_g} \tag{18}$$

and $\boldsymbol{\chi}_k$ denotes a $n_g \times n_g$ matrix with (i, j)-entry χ_{kij}.

The foregoing stationary model for asynchronous data with imperfect coherence has been investigated and verified in a previous paper using synthetic and experimental data by the authors [39]. The validation results based on an experimental data example in the paper are summarised here. Two accelerometers were placed on a one-storey shear frame structure with the same measured DOFs. The analog data from these two sensors were sampled using two different DAQ units. The DAQ units were controlled to start and finish measuring at the same time while using their own clock for sampling. Two hours of data from these two sensors were collected with the sampling rate of 2048Hz. Figure 1 shows the root singular value spectrum of the measured data (solid lines) and the simulated values based on the proposed asynchronous data model (dashed lines). It can be seen that the theoretical predictions agree well with the values calculated from real asynchronous data, verifying the relevance of the proposed model. Multiple trials of data have been investigated (results omitted here), showing qualitatively similar agreement.

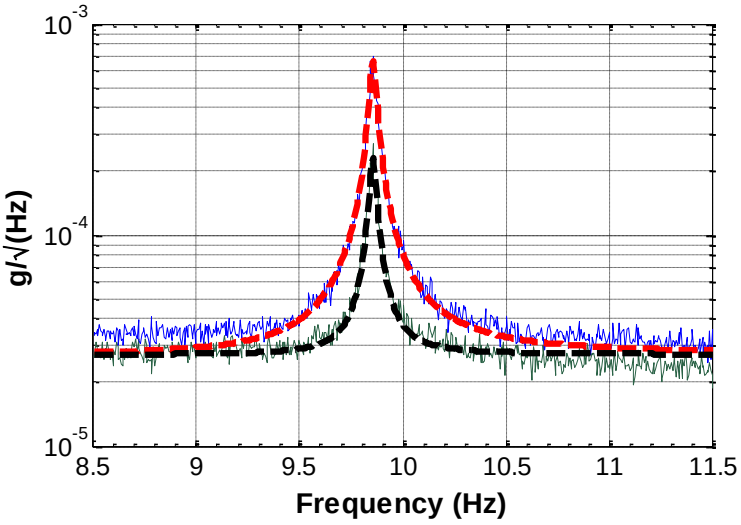


Figure 1. Root Singular Value Spectrum of the Validation Example (Solid Line-Exact Values, Dashed Line-Simulated Values)

4. Mathematical Structure of PSD Matrix

Recall that the MPV of modal parameters minimises the NLLF in Eq.(5). As is typical in developing efficient algorithms for MPV, it is necessary to express the determinant and inverse of $\mathbf{E}_k$ in better forms to facilitate computations without resorting to brute-force numerical optimisation.

For instructional purpose, first consider the case when data is synchronised, where $\mathbf{E}_k$ is given by Eq.(10). The standard technique is to consider an orthonormal basis $\{\mathbf{a}_j \in R^n\}_{j=1}^n$ with $\mathbf{a}_1 = \boldsymbol{\varphi} \in R^n$ (assuming $\|\boldsymbol{\varphi}\|^2 = 1$) and then write $\mathbf{I}_n = \sum_{j=1}^n \mathbf{a}_j \mathbf{a}_j^T$. Substituting into Eq.(10) gives the eigenvector representation of $\mathbf{E}_k$:

$$\mathbf{E}_k = \underbrace{(SD_k + S_e)}_{Eigenvalue} \mathbf{a}_1 \mathbf{a}_1^T + \sum_{j=2}^{n} \underbrace{S_e}_{Eigenvalue} \mathbf{a}_j \mathbf{a}_j^T \tag{19}$$

The determinant of $\mathbf{E}_k$ is simply the product of the eigenvalues, i.e.,

$$\det \mathbf{E}_k = (SD_k + S_e) S_e^{n-1} \tag{20}$$

The inverse of $\mathbf{E}_k$ has the same eigenvectors but reciprocal of eigenvalues:

$$\mathbf{E}_k^{-1} = (SD_k + S_e)^{-1} \mathbf{a}_1 \mathbf{a}_1^T + \sum_{j=2}^{n} S_e^{-1} \mathbf{a}_j \mathbf{a}_j^T \tag{21}$$

Back-substituting $\sum_{j=2}^n \mathbf{a}_j \mathbf{a}_j^T = \mathbf{I}_n - \boldsymbol{\varphi}\boldsymbol{\varphi}^T$ eliminates $\{\mathbf{a}_j\}_{j=2}^n$:

$$\mathbf{E}_k^{-1} = S_e^{-1} \mathbf{I}_n - S_e^{-1} (1 + S_e / SD_k)^{-1} \boldsymbol{\varphi}\boldsymbol{\varphi}^T \tag{22}$$

The situation is more complicated for asynchronous data when $\mathbf{E}_k$ is given by Eq.(17). For simplicity, assume the same prediction error S_e for different groups, so that

$$\mathbf{E}_k = SD_k \mathbf{U} \mathbf{C}_k \mathbf{U}^T + S_e \mathbf{I}_n \tag{23}$$

Let $\{\lambda_{ki} \geq 0\}_{i=1}^{n_g}$ and $\{\mathbf{c}_{ki} \in C^{n_g}\}_{i=1}^{n_g}$ be the eigenvalues and eigenvectors (with unit norm) of $\mathbf{C}_k$. Then

$$\mathbf{C}_k = \sum_{i=1}^{n_g} \lambda_{ki} \mathbf{c}_{ki} \mathbf{c}_{ki}^* \tag{24}$$

Substituting Eq.(24) into Eq.(23) gives:

$$\mathbf{E}_k = \sum_{i=1}^{n_g} SD_k \lambda_{ki} \boldsymbol{\alpha}_{ki} \boldsymbol{\alpha}_{ki}^* + S_e \mathbf{I}_n \tag{25}$$

where

$$\boldsymbol{\alpha}_{ki} = \mathbf{U}\mathbf{c}_{ki} \qquad i = 1,..., n_g \tag{26}$$

Using $\mathbf{U}^T\mathbf{U} = \mathbf{I}_{n_g}$ and the orthonormal properties of $\{\mathbf{c}_{ki}\}_{i=1}^{n_g}$, i.e., $\mathbf{c}_{ki}^*\mathbf{c}_{ki} = 1$ and $\mathbf{c}_{ki}^*\mathbf{c}_{kj} = 0$ $(i \neq j)$, it can be verified that $\{\boldsymbol{\alpha}_{ki}\}_{i=1}^{n_g}$ form an orthonormal basis in a n_g-dimensional subspace of C^n. Let $\{\mathbf{a}_{ki}\}_{i=n_g+1}^{n}$ be an orthonormal basis in the orthogonal complement of this subspace. Also, let $\mathbf{a}_{ki} = \boldsymbol{\alpha}_{ki}$ ($i = 1,..., n_g$). Then $\{\mathbf{a}_{ki}\}_{i=1}^{n}$ is an orthonormal basis in C^n. Substituting $\mathbf{I}_n = \sum_{i=1}^{n} \mathbf{a}_{ki}\mathbf{a}_{ki}^*$ into Eq.(25) gives the eigenvector representation of $\mathbf{E}_k$:

$$\mathbf{E}_k = \sum_{i=1}^{n_g} \underbrace{\left(SD_k \lambda_{ki} + S_e\right)}_{Eigenvalue} \mathbf{a}_{ki}\mathbf{a}_{ki}^* + \sum_{i=n_g+1}^{n} \underbrace{S_e}_{Eigenvalue} \mathbf{a}_{ki}\mathbf{a}_{ki}^* \tag{27}$$

It follows that

$$\det \mathbf{E}_k = S_e^{n-n_g} \prod_{i=1}^{n_g} \left(SD_k \lambda_{ki} + S_e\right) \tag{28}$$

and

$$\mathbf{E}_k^{-1} = \sum_{i=1}^{n_g} \left(SD_k \lambda_{ki} + S_e\right)^{-1} \mathbf{a}_{ki}\mathbf{a}_{ki}^* + \sum_{i=n_g+1}^{n} S_e^{-1} \mathbf{a}_{ki}\mathbf{a}_{ki}^* \tag{29}$$

Back-substituting $\sum_{i=n_g+1}^{n} \mathbf{a}_{ki}\mathbf{a}_{ki}^* = \mathbf{I}_n - \sum_{i=1}^{n_g} \mathbf{U}\mathbf{c}_{ki}\mathbf{c}_{ki}^*\mathbf{U}^*$ into Eq.(29) eliminates $\{\mathbf{a}_{ki}\}_{i=1}^{n}$:

$$\mathbf{E}_k^{-1} = S_e^{-1}\mathbf{I}_n - \sum_{i=1}^{n_g} S_e^{-1}\left(1 + S_e / SD_k \lambda_{ki}\right)^{-1} \mathbf{U}\mathbf{c}_{ki}\mathbf{c}_{ki}^*\mathbf{U}^* \tag{30}$$

The foregoing results illustrate the differences between synchronous and asynchronous data in their singular value spectrum (i.e. the plot of eigenvalues of the PSD matrix), which is commonly used for detecting potential modes in OMA. In Eq.(19), there is only one significant eigenvalue $SD_k + S_e$ for synchronous data around the resonance band of a well-

separated mode. The remaining ones are all equal to S_e, which indicate the noise level. For asynchronous data, however, there are n_g significant eigenvalues $\{SD_k\lambda_{ki} + S_e\}_{i=1}^{n_g}$ with the common dynamic amplification factor D_k (see Eq.(27)). This leads to additional peaks with the same variation with frequency in the singular value spectrum for each mode, which appear as several extremely close modes with almost identical damping ratios.

In the present case, the determinant and inverse of $\mathbf{E}_k$ depend on the eigenvalue λ_{ki} and eigenvector $\mathbf{c}_{ki}$ of $\mathbf{C}_k$. The latter depend in a non-trivial manner on the local mode shapes $\{\|\mathbf{u}_i\|\}_{i=1}^{n_g}$ as well as the coherence $\boldsymbol{\chi}_k$ (see Eq. (18)). The eigenvalue properties will become even more complicated in the general case when the prediction error may differ for different synchronous groups.

5. Zero Coherence Formulation

As shown in the previous section, it is difficult to express the determinant and inverse of the PSD matrix $\mathbf{E}_k$ for asynchronous data in analytically tractable form. To facilitate the efficient determination of MPV of modal parameters, an algorithm is developed that assumes zero coherence among different groups. This is justified for small coherence. The effect of this approximation on the MPV and posterior uncertainty shall be investigated through numerical examples in the companion paper. Based on the zero coherence assumption, analytical expressions of the determinant and inverse of $\mathbf{E}_k$ can be obtained. The resulting form of the NLLF is derived in this section.

Assuming that the coherences between different synchronous groups are zero, $\boldsymbol{\chi}_k$ now becomes an identity matrix and the theoretical PSD matrix $\mathbf{E}_k$ in Eq.(17) has a block-diagonal form:

$$\mathbf{E}_k = \begin{bmatrix} \mathbf{E}_{1k} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_{2k} & & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{E}_{n_g k} \end{bmatrix} \tag{31}$$

where

$$\mathbf{E}_{ik} = SD_k c_i \bar{\mathbf{u}}_i \bar{\mathbf{u}}_i^T + S_{ei}\mathbf{I}_{n_i} \tag{32}$$

and

$$\bar{\mathbf{u}}_i = \frac{\mathbf{u}_i}{\|\mathbf{u}_i\|} \tag{33}$$

$$c_i = \|\mathbf{u}_i\|^2 \tag{34}$$

The determinant and inverse of $\mathbf{E}_k$ are given by

$$\det(\mathbf{E}_k) = \prod_{i=1}^{n_g} \det(\mathbf{E}_{ik}) \tag{35}$$

$$\mathbf{E}_k^{-1} = \begin{bmatrix} \mathbf{E}_{1k}^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_{2k}^{-1} & & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{E}_{n_g k}^{-1} \end{bmatrix} \tag{36}$$

Substituting Eq.(35) and Eq.(36) into the NLLF in Eq.(5), the resulting NLLF now can be written as:

$$L = \sum_{i=1}^{n_g} L_i \tag{37}$$

with

$$L_i = \sum_k \ln\det \mathbf{E}_{ik} + \sum_k \mathcal{F}_{ik}^{*} \mathbf{E}_{ik}^{-1} \mathcal{F}_{ik} \tag{38}$$

where $\mathcal{F}_{ik}$ is the corresponding partition of $\mathcal{F}_k$ for the i th synchronous group.

As $\mathbf{E}_{ik}$ has a similar form of $\mathbf{E}_k$ for synchronous data, Eq.(20) and Eq.(22) can be applied to give

$$\det(\mathbf{E}_{ik}) = (SD_k c_i + S_{ei}) S_{ei}^{n_i - 1} \tag{39}$$

and

$$\mathbf{E}_{ik}^{-1} = S_{ei}^{-1}\mathbf{I}_n - S_{ei}^{-1}(1 + S_{ei} / SD_k c_i)^{-1} \bar{\mathbf{u}}_i \bar{\mathbf{u}}_i^T \tag{40}$$

Substituting Eq.(39) and Eq.(40) into Eq.(38) and writing $(\mathcal{F}_{ik}^{*}\bar{\mathbf{u}}_i)(\bar{\mathbf{u}}_i^T \mathcal{F}_{ik}) = (\bar{\mathbf{u}}_i^T \mathcal{F}_{ik})(\mathcal{F}_{ik}^{*}\bar{\mathbf{u}}_i)$ (since $\mathcal{F}_{ik}^{*}\bar{\mathbf{u}}_i$ is a scalar) give the resulting NLLF as:

$$L_i = (n_i - 1)N_f \ln S_{ei} + \sum_k \ln\left(SD_k c_i + S_{ei}\right) + S_{ei}^{-1}\left(d_i - \overline{\mathbf{u}}_i^T \mathbf{A}_i \overline{\mathbf{u}}_i\right) \tag{41}$$

where

$$d_i = \sum_k \mathcal{F}_{ik}^{\ *} \mathcal{F}_{ik} \tag{42}$$

and

$$\mathbf{A}_i = \sum_k \left(1 + S_{ei} / SD_k c_i\right)^{-1} \mathbf{D}_{ik} \tag{43}$$

$$\mathbf{D}_{ik} = \mathcal{F}_{ik} \mathcal{F}_{ik}^{\ *} \tag{44}$$

6. Fast Algorithm for MPV

Based on the zero coherence formulation, an iterative scheme is proposed in this section that allows the MPV of modal parameters to be determined efficiently. Because of the nonlinear nature of the NLLF, generally it is not possible to obtain the MPV analytically. The modal parameters to be identified are first re-parameterised to facilitate optimisation. As the NLLF is a quadratic form in the local mode shape for each synchronous group, the MPV of the local mode shapes can be obtained analytically in terms of other parameters, and vice-versa. This leads to an iterative procedure to be proposed. The computational cost of this procedure is insensitive to the measured DOFs n as the MPV of the local mode shapes can be obtained by solving an eigenvalue problem. The resulting modal identification algorithm effectively suppresses the growth of computational effort with the increase of n, which provides a fast estimation of the modal parameters even for large scale tests.

According to the NLLF developed in the last section, the set of parameters to be identified includes $\{f, \zeta, S, \{S_{ei}\}, \{c_i\}, \{\overline{\mathbf{u}}_i\}\}$ with the constraints

$$\sum_{i=1}^{n_g} c_i = 1 \tag{45}$$

and

$$\left\|\overline{\mathbf{u}}_i\right\| = 1 \qquad\qquad i = 1, \ldots, n_g \tag{46}$$

Evaluating $\{c_i\}$ directly with the constraints in Eq.(45) involves solving nonlinear equations. To see this, the Lagrangian for this constrained optimisation problem is given by

$$J = \sum_{i=1}^{n_g} \left[(n_i - 1) N_f \ln S_{ei} + \sum_k \ln\left(SD_k c_i + S_{ei}\right) + S_{ei}^{-1}\left(d_i - \bar{\mathbf{u}}_i^T \mathbf{A}_i \bar{\mathbf{u}}_i\right) \right] + q\left(\sum_{i=1}^{n_g} c_i - 1 \right) \tag{47}$$

where q is the Lagrange multiplier that enforces the norm of the global mode shape to be 1 (or equivalently Eq.(45)). The gradient of J with respect to $\{c_i\}$ is given by

$$\frac{dJ}{dc_i} = \sum_k \left(c_i + \frac{S_{ei}}{SD_k} \right)^{-1} - S_{ei}^{-1} \bar{\mathbf{u}}_i^T \frac{d\mathbf{A}_i}{dc_i} \bar{\mathbf{u}}_i + q \tag{48}$$

where

$$\frac{d\mathbf{A}_i}{dc_i} = \sum_k \frac{S_{ei}}{SD_k} c_i^{-2} \left(1 + S_{ei} / SD_k c_i\right)^{-2} \mathbf{D}_{ik} \tag{49}$$

The MPV of $\{c_i\}$ should be obtained by solving $dJ / dc_i = 0$ using Eq.(48). Generally, it is not possible to obtain the analytical form of c_i in terms of other parameters. In view of this, combine S and $\{c_i\}$ under the constraints Eq.(45) into free parameters. Let

$$S_i = S c_i \qquad\qquad i = 1, \ldots, n_g \tag{50}$$

The parameter set $\boldsymbol{\theta}$ now becomes $\{f, \zeta, \{S_i\}, \{S_{ei}\}, \{\bar{\mathbf{u}}_i\}\}$ under constraints Eq.(46). In terms of these parameters,

$$L_i = (n_i - 1) N_f \ln S_{ei} + \sum_k \ln\left(S_i D_k + S_{ei}\right) + S_{ei}^{-1}\left(d_i - \bar{\mathbf{u}}_i^T \mathbf{A}_i \bar{\mathbf{u}}_i\right) \tag{51}$$

where

$$\mathbf{A}_i = \sum_k \left(1 + S_{ei} / S_i D_k\right)^{-1} \mathbf{D}_{ik} \tag{52}$$

Thus the MPV of $\{S_i\}$ are first determined by unconstrained optimisation, from which the MPV of S and $\{c_i\}$ can be recovered. Note that f and ζ are common to all L_i. They are the only parameters that connect the NLLF of different groups. For each group i, the MPV of S_i, S_{ei} and $\bar{\mathbf{u}}_i$ can be determined by minimizing L_i only. To obtain the MPV of $\{\bar{\mathbf{u}}_i\}$, the term $S_{ei}^{-1}\left(d_i - \bar{\mathbf{u}}_i^T \mathbf{A}_i \bar{\mathbf{u}}_i\right)$ in the NLLF (see Eq.(51)) should be minimised subjected to constraint in Eq.(46). This can be done by introducing the constraint into Eq. (51) using Lagrange multiplier and setting the first derivative of the resulting function with respect to $\bar{\mathbf{u}}_i$ equal to zero. It follows that the MPV of $\bar{\mathbf{u}}_i$ is the eigenvector of $\mathbf{A}_i$ in Eq.(52) with the largest eigenvalue.

As the coherences are assumed to be zero, the relative sense of the mode shapes between different groups cannot be identified. That is, $+\bar{\mathbf{u}}_i$ and $-\bar{\mathbf{u}}_i$ give the same value in the NLLF. In practice, the relative sense can be determined based on intuition, e.g., spatial continuity of mode shapes. Of course, for complicated modes (e.g., high frequency modes) this may not be easy. This is one fundamental limitation of zero coherence data.

7. High Signal-to-noise Asymptotics

The asymptotic behaviour of the MPV of modal parameters for high signal-to-noise ratio (SNR) data is studied in this section. It provides a quick estimation of some modal parameters, which can also be used as initial guesses for the iteration procedure presented later.

Motivated by the form of $\mathbf{A}_i$ in Eq.(43), high SNR here refers to

$$\gamma_{ki} = S_i D_k / S_{ei} = S D_k c_i / S_{ei} >> 1 \qquad i = 1,...,n_g \tag{53}$$

This is the SNR of a single setup scaled by $c_i = \|\mathbf{u}_i\|^2$. As a result of Eq.(53),

$$\left[1 + S_{ei} / S_i D_k\right]^{-1} = \left[1 + \gamma_{ki}^{-1}\right]^{-1} \sim 1 - \gamma_{ki}^{-1} \sim 1 \tag{54}$$

Consider the zeroth order approximation of $\mathbf{A}_i$:

$$\mathbf{A}_i \sim \sum_{k=1}^{N_f} \mathcal{F}_{ik} \mathcal{F}_{ik}^{*} = \mathbf{A}_{0i} \tag{55}$$

which becomes a constant matrix. This means that the high SNR MPV of $\bar{\mathbf{u}}_i$ can be obtained directly as the eigenvector of $\mathbf{A}_{0i}$ with the largest eigenvalue and scaled to have unit norm.

The high SNR MPV of S_i and S_{ei} can be obtained using the first order approximation of $\mathbf{A}_i$:

$$\mathbf{A}_i \sim \sum_{k=1}^{N_f} \left[1 - S_{ei} / S_i D_k\right] \mathcal{F}_{ik} \mathcal{F}_{ik}^{*} = \mathbf{A}_{0i} - S_{ei} S_i^{-1} \sum_{k=1}^{N_f} D_k^{-1} \mathbf{D}_{ik} \tag{56}$$

Substituting Eq.(56) to Eq.(51) gives:

$$L_i = \sum_{k=1}^{N_f} \ln D_k + (n_i - 1)N_f \left\{ \ln S_{ei} + S_{ei}^{-1} \left[\left(d_i - \bar{\mathbf{u}}_i^T \mathbf{A}_{0i} \bar{\mathbf{u}}_i \right) / (n_i - 1)N_f \right] \right\} + N_f \left\{ \ln S_i + (S_i)^{-1} \left[N_f^{-1} \sum_{k=1}^{N_f} \bar{\mathbf{u}}_i^T D_k^{-1} \mathbf{D}_{ik} \bar{\mathbf{u}}_i \right] \right\} \quad (57)$$

This expression shows that L_i depends on S_i and S_{ei} through the first and second braces, which are of the form $\ln x + c/x$. This form has a unique minimum of $1 + \ln c$ at $x = c$. Considering the first brace, the high SNR asymptotic MPV of S_{ei} is given by

$$\hat{S}_{ei} \sim \left(d - \hat{\lambda}_{0i} \right) / (n_i - 1)N_f \quad (58)$$

where

$$\hat{\lambda}_{0i} = \bar{\mathbf{u}}_i^T \mathbf{A}_{0i} \bar{\mathbf{u}}_i \quad (59)$$

is the maximum eigenvalue of $\mathbf{A}_{0i}$. Minimizing the second brace with respect to S_i gives the high SNR asymptotic MPV of S_i:

$$\hat{S}_i \sim N_f^{-1} \sum_{k=1}^{N_f} \bar{\mathbf{u}}_i^T D_k^{-1} \mathbf{D}_{ik} \bar{\mathbf{u}}_i \quad (60)$$

8. Summary of Procedure

Based on the foregoing analysis, an iterative procedure for determining the MPV of modal parameters is summarised as follow.

Step I. Initial Guess

1) Calculate the FFT of the measured data and plot the singular value spectrum.
2) Select the frequency band for the mode of interest.
3) Obtain the initial guess of f from the singular value spectrum and set the initial guess of ζ as 1%.
4) Take the initial guess of $\bar{\mathbf{u}}_i$ $(i = 1, ..., n_g)$ as the eigenvector with largest eigenvalue of $\mathbf{A}_{0i}$ in Eq.(55).
5) Calculate the initial guess of S_i and S_{ei} $(i = 1, ..., n_g)$ using Eq.(60) and Eq.(58) respectively.

Step II. Iteration Phase

In the following, parameters that are not updated are kept at their current value during iteration.

1) Update $\{f, \zeta\}$ by minimizing L in Eq.(37) with respect to $\{f, \zeta\}$.
2) Update S_i and S_{ei} $(i = 1,..., n_g)$ by minimizing L_i in Eq.(51).
3) Update $\bar{\mathbf{u}}_i$ $(i = 1,..., n_g)$ as the eigenvector of $\mathbf{A}_i$ in Eq.(52) with the largest eigenvalue.

Repeat Steps 1) to 3) until convergence.

Step III. MPV of S and $\{c_i\}$

Obtain the MPV of S by

$$\hat{S} = \sum_{i=1}^{n_g} \hat{S}_i \tag{61}$$

For $i = 1,..., n_g$, obtain the MPV of c_i by

$$\hat{c}_i = \frac{\hat{S}_i}{\hat{S}} \tag{62}$$

Step IV. Global Mode Shape

Obtain the MPV of global mode shape by

$$\hat{\boldsymbol{\varphi}} = \begin{bmatrix} s_1 \sqrt{\hat{c}_1}\, \hat{\bar{\mathbf{u}}}_1 \\ \vdots \\ s_i \sqrt{\hat{c}_i}\, \hat{\bar{\mathbf{u}}}_i \\ \vdots \\ s_{n_g} \sqrt{\hat{c}_{n_g}}\, \hat{\bar{\mathbf{u}}}_{n_g} \end{bmatrix} \tag{63}$$

where $s_i = \pm 1$ is determined by plotting the mode shape to fit intuition.

9. Illustrative Examples

Two examples are presented to illustrate the proposed method. The first example serves to verify consistency based on synthetic data, where there is no modelling error. The second example is based on laboratory data, which investigates applicability to real asynchronous data while still under well controlled environment. This section focuses on MPVs.

Identification uncertainty is investigated in the companion paper, where an application to field data is also presented.

9.1 Synthetic Data

Consider a six-storey shear building as shown in Figure 2. It has a uniform inter-storey stiffness of 3000 kN/mm and floor mass of 600 tons. The natural frequency of the first mode is calculated to be 2.71Hz. The damping ratios of the first four modes are assumed to be 1%, 1%, 1.3% and 1.5%, respectively. The structure is subjected to independent and identically distributed (i.i.d.) Gaussian white noise excitation with a PSD of $11.8\text{N}/\sqrt{\text{Hz}}$ in the horizontal direction of every floor. The resulting acceleration response is in the order of few tens of $\mu\text{g}/\sqrt{\text{Hz}}$ around the resonance peaks of modes, which is typical in ambient field tests. The measured acceleration data comprises three synchronous groups as shown in Figure 2. To simulate the asynchronous data, the coherences of the modal excitations among different groups are taken to be zero (i.e., modal excitations are generated independently). The acceleration response of the structure is generated at a sampling rate of 100Hz with a duration of 1000s. The acceleration data is contaminated by Gaussian white noise with PSD of a few $\mu\text{g}/\sqrt{\text{Hz}}$, which simulates the typical noise level of force-balanced accelerometers. The noise levels are $2\mu\text{g}/\sqrt{\text{Hz}}$, $3\mu\text{g}/\sqrt{\text{Hz}}$ and $5\mu\text{g}/\sqrt{\text{Hz}}$ for Groups 1 to 3, respectively.

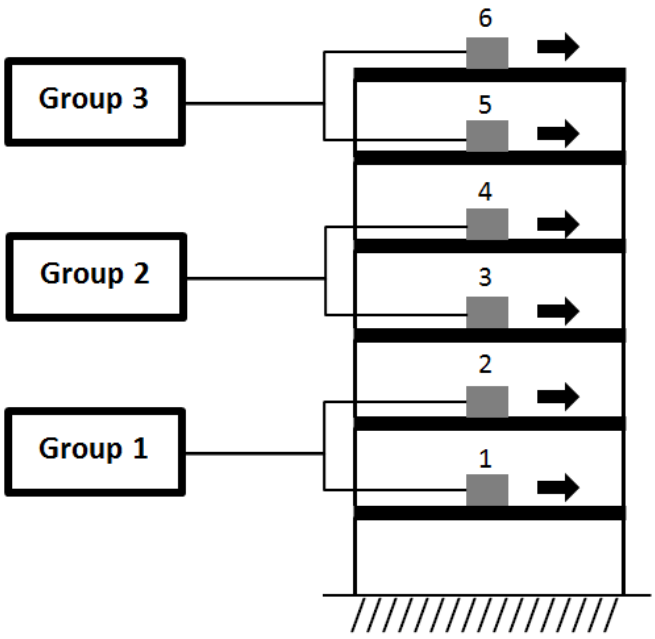


Figure 2. Schematic Diagram, Synthetic Data Example

Figure 3 shows the root singular value spectrum of the data. The first four modes are identified separately using the proposed method. The (hand-picked) initial guesses and selected frequency bands are shown with a circle and the symbol '[-]', respectively. The MPV of the modal parameters and their exact values that generated the data are summarised in Table 1. The MPVs generally agree well with their exact values. As the mode number increases, the MPV of prediction error PSD tends to be larger than the assumed measurement noise. This is due to unexplained contribution from the non-resonant part of the lower modes (i.e. modelling error). The identified mode shapes (denoted by squares) and the exact values (denoted by solid lines) are shown in Figure 4. They agree well with the exact ones. The MAC (Modal Assurance Criteria) between the identified and exact mode shape is calculated to be 0.9999, 0.9998, 0.9997, 0.9999 for Mode 1 to 4, respectively.

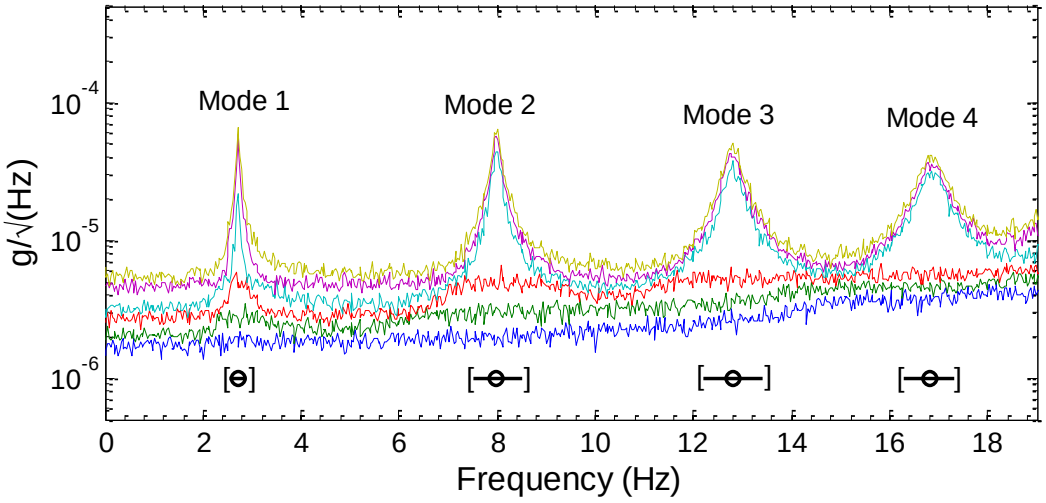


Figure 3. Root Singular Value Spectrum, Synthetic Data Example

Table 1. Identified Modal Parameters (MPVs), Synthetic Data Example

Mode	f (Hz)		ζ (%)		$\sqrt{S}$ (μg/$\sqrt{\text{Hz}}$)		$\sqrt{S_{ej}}$ (μg/$\sqrt{\text{Hz}}$)	
	MPV	Exact	MPV	Exact	MPV	Exact	MPV	Exact
1	2.710	2.713	0.95	1.00	1.96	2.00	2.06 2.85 5.15	2.00 3.00 5.00
2	7.980	7.981	1.02	1.00	2.00	2.00	2.07	2.00

							3.13	3.00
							5.08	5.00
3	12.792	12.786	1.30	1.30	1.91	2.00	2.76	2.00
							3.62	3.00
							5.44	5.00
4	16.849	16.847	1.50	1.50	1.83	2.00	4.19	2.00
							4.67	3.00
							5.63	5.00

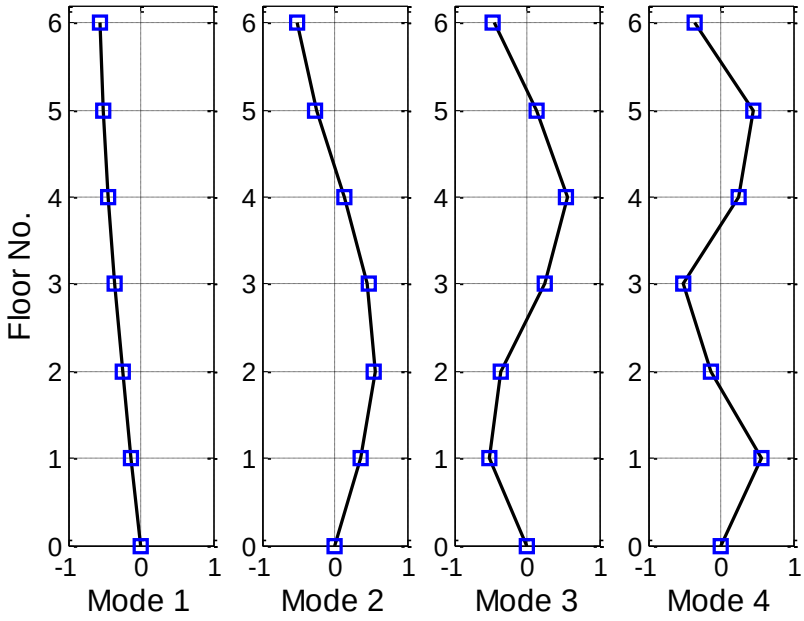


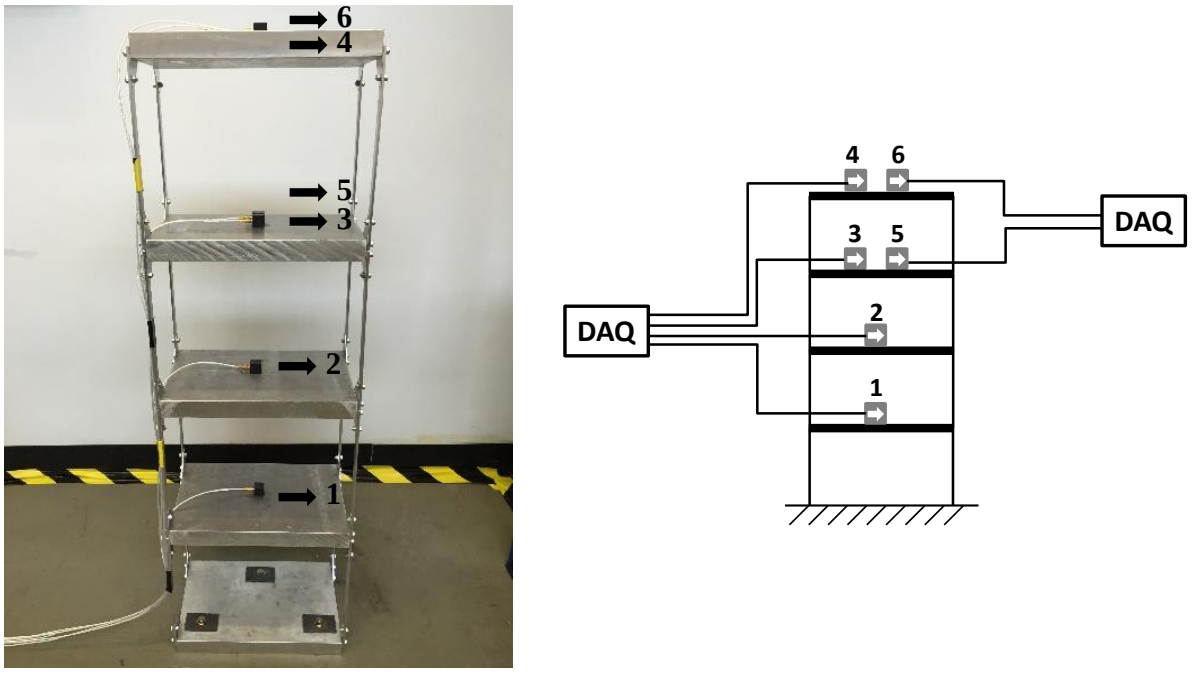
Figure 4. Mode Shapes, Synthetic Data Example (Solid Line-Exact, Square-MPV)

9.2 Laboratory Shear Building Model

Consider a four-storey laboratory shear frame as shown in Figure 5, measuring $30\text{cm} \times 20\text{cm}$ in plan with a uniform storey height of 25cm. Six piezoelectric accelerometers distributed at the centre of each floor are used to measure the vibration in the weak direction (parallel to paper). Sensors 1 to 4 are synchronised using one DAQ unit and Sensors 5 & 6 are synchronised using another. Data recording of all data channels was set to start and finish at the same time with the same sampling rate.

Thirty minutes of ambient data was sampled at 2048Hz and later decimated to 256Hz for analysis. The data comprising channel 1 to 4 is referred as the 'synchronous data set'. The

MPV of modal parameters based on this set of data are identified using the fast Bayesian FFT method for synchronous data [40]. The identification results are used as the reference values. The data comprising channel $\{1,2,5,6\}$ is referred as the 'asynchronous data set' where the MPV of modal parameters are identified using the proposed method.



(a) (b)

Figure 5. Four-storey Laboratory Shear Building (a) Side View (b) Setup

The root singular value spectrum for the asynchronous data set is plotted in Figure 6. Modal identification here focuses on the first five modes. For each mode there are two lines significantly above the remaining ones, which indicate imperfect synchronisation. This is also consistent with the fact that there are two synchronous groups. Table 2 shows the identified MPVs of the modal parameters for both synchronous data set and asynchronous data set. The MPV of modal parameters determined by the proposed method using the asynchronous data set are very close to the ones identified based on the synchronous data set. The identified mode shape MPVs are shown in Figure 7 for both the synchronous (solid lines) and asynchronous (squares) data sets. They almost coincide. The MAC values between these two types are calculated to be 0.9995, 0.9998, 0.9975, 0.9998 and 0.9999 from Mode 1 to 5, respectively. The identification results illustrate that the proposed method can also provide a good estimation of the modal parameters when dealing with real asynchronous data.

For reference, conventional methods that assume synchronous data have also been applied to modal identification of the asynchronous data set. The identified natural frequencies and damping ratios are close to the ones identified based on synchronous data (details omitted here). However, the relative scaling and directions between the identified local mode shapes of different synchronous groups are erroneously determined. Figure 8 shows the identified mode shape for the synchronous data sets (solid lines) and asynchronous data set (circles) using the fast Bayesian FFT method that assumes synchronous data [40]. It can be seen that for Mode 2 to Mode 5, the relative direction of the identified partial mode shapes on third and forth floor are incorrect. There are also some errors in the identified mode shape values on the first and second floor for Mode 1 and Mode 2.

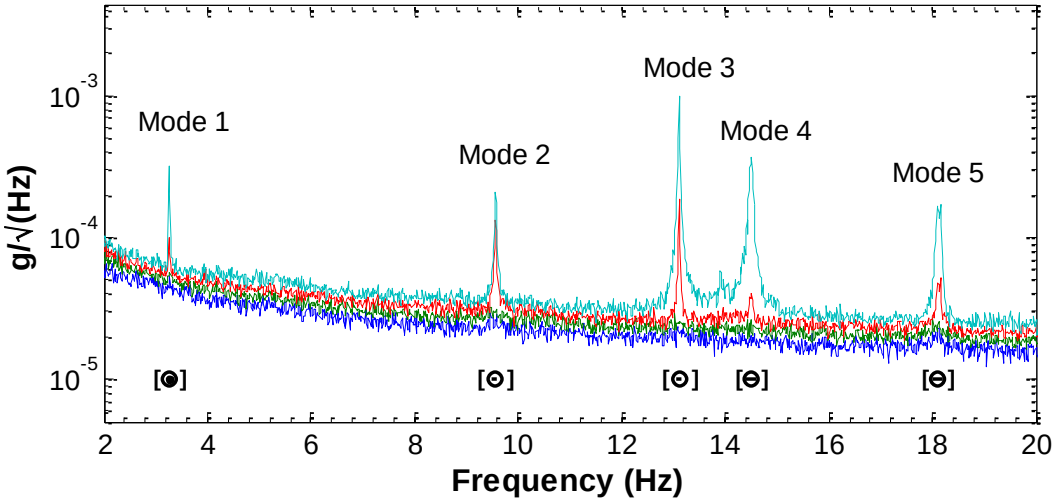


Figure 6. Root Singular Value Spectrum for Asynchronous Data Set, Laboratory Shear Building

Table 2. Identified Modal Parameters (MPVs), Laboratory Shear Building

Mode	f (Hz)		ζ (%)		$\sqrt{S}$ (μg/$\sqrt{\text{Hz}}$)		$\sqrt{S_{ej}}$ (μg/$\sqrt{\text{Hz}}$)	
	Asyn.	Syn.	Asyn.	Syn.	Asyn.	Syn.	Asyn.	Syn.
1	3.260	3.260	0.08	0.07	0.97	0.89	54.1 50.1	54.1
2	9.560	9.560	0.15	0.15	0.74	0.76	28.9 27.8	28.1
3	13.108	13.108	0.04	0.04	1.05	1.03	22.9	25.1

							23.8	
4	14.496	14.496	0.15	0.15	1.21	1.21	20.5 22.9	23.5
5	18.128	18.128	0.23	0.22	0.83	0.83	22.7 21.1	22.0

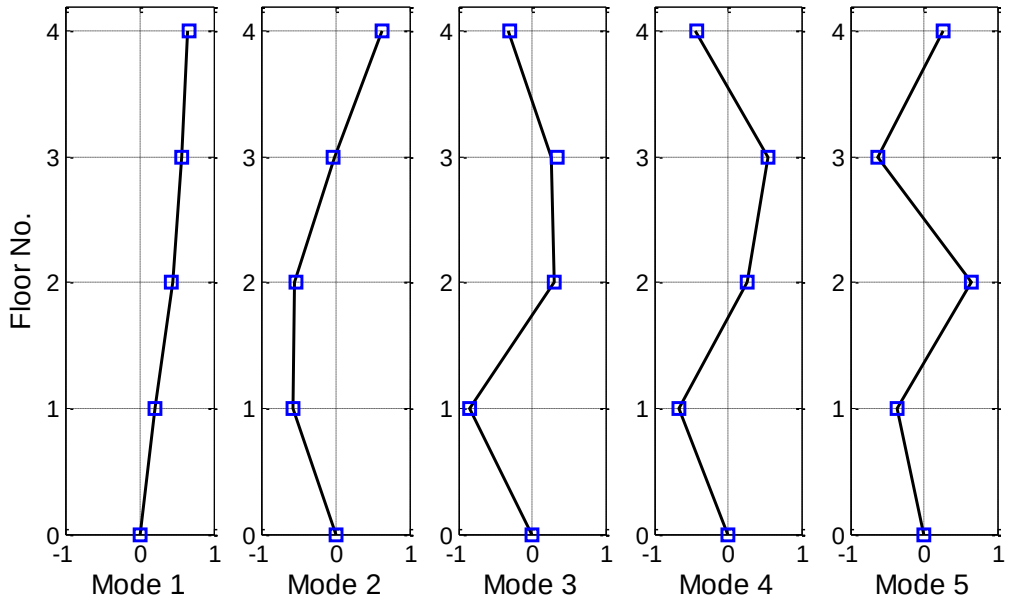


Figure 7. Identified Mode Shape MPVs (Solid Line-Synchronous Data Set, Square-Asynchronous Data Set)

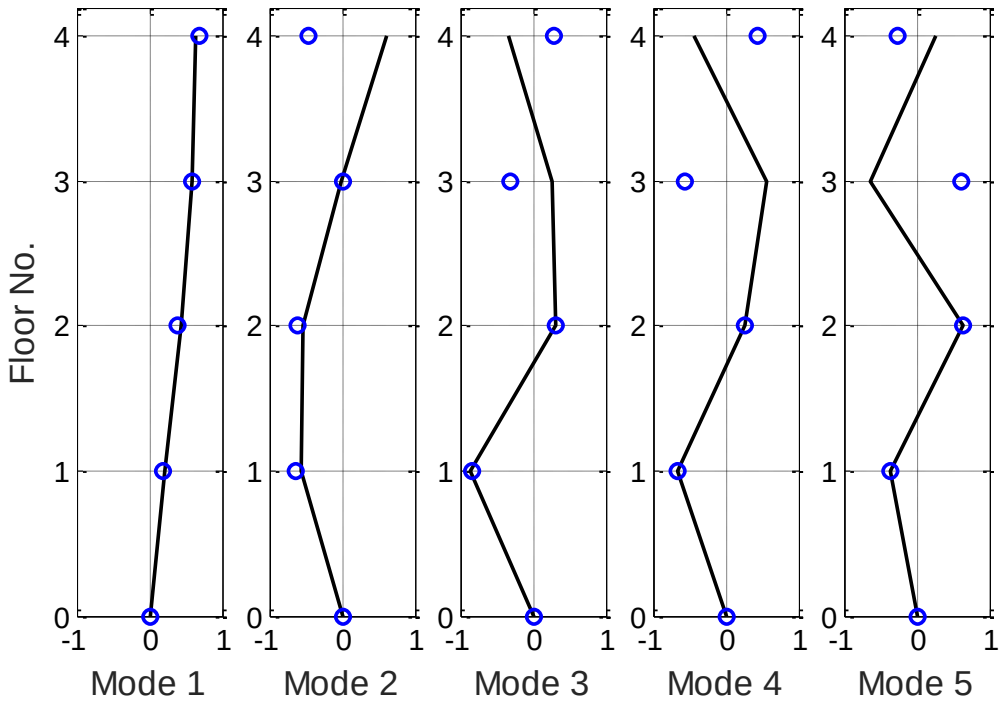


Figure 8. Identified Mode Shape MPVs (Solid Line-Synchronous Data Set, Circle-Asynchronous Data Set using Conventional Method)

10 Conclusions

A fast Bayesian OMA formulation for asynchronous data has been proposed in this work. Asynchronous data is generally a non-stationary process that is difficult to model or analyse from first principles. A stationary stochastic model with imperfect coherence has been proposed, which is conducive to analysis while retaining the characteristics of asynchronous nature over suitable time scales. The likelihood function depends in a complicated manner on the coherence parameters, making it difficult to develop fast algorithms. In view of this, an alternative formulation is proposed that assumes zero coherence. Significant simplifications result, based on which a fast algorithm has been developed for determining the most probable value (MPV) of modal parameters.

This paper has focused on the determination of MPV, whose consistency has been demonstrated in the synthetic and laboratory data examples. In the companion paper (Part II), an efficient method for determining the identification uncertainty shall be developed. The proposed method is further verified with a parametric study to investigate the effect of

modelling error due to zero coherence assumption. An application with field data shall also be presented.

Acknowledgments

This paper was partly supported by Tung Doctoral Scholarship and UK Engineering & Physical Sciences Research Council (EP/N017897/1). The financial support is gratefully acknowledged.

References

[1] D.E. Hudson, Dynamic Tests of Full-Scale Structures, J. Eng. Mech. Div. 103 (1977) 1141–1157.

[2] D.J. Ewins, Modal testing : theory, practice, and application, Baldock : Research Studies Press, 2000., 2000.

[3] S. Beskhyroun, L.D. Wegner, B.F. Sparling, New methodology for the application of vibration-based damage detection techniques, Struct. Control Heal. Monit. (2011). doi:10.1002/stc.

[4] G. Hearn, R. Testa, Modal analysis for damage detection in structures, J. Struct. Eng. 117 (1991) 3042–3063.

[5] C.R. Farrar, K. Worden, An introduction to structural health monitoring., Philos. Trans. A. Math. Phys. Eng. Sci. 365 (2007) 303–315. doi:10.1098/rsta.2006.1928.

[6] F.N. Catbas, T. Kijewski-Correa, A.E. Aktan, Structural identification (St-Id) of constructed facilities: Approaches, methods and technologies for effective practice of St-Id, in: Am Soc Civ Eng, 2011.

[7] S.S. Ivanović, M.D. Trifunac, E.I. Novikova, a. a. Gladkov, M.I. Todorovska, Ambient vibration tests of a seven-story reinforced concrete building in Van Nuys, California, damaged by the 1994 northridge earthquake, Soil Dyn. Earthq. Eng. 19 (2000) 391–411. doi:10.1016/S0267-7261(00)00025-7.

[8] H. Wenzel, D. Pichler, Ambient vibration monitoring, Wiley, UK, 2005.

[9] J.M.W. Brownjohn, Structural health monitoring of civil infrastructure., Philos. Trans. A. Math. Phys. Eng. Sci. 365 (2007) 589–622. doi:10.1098/rsta.2006.1925.

[10] C. Rainieri, G. Fabbrocino, Operational Modal Analysis of Civil Engineering Structures, Springer, New York, NY, 2014. doi:10.1007/978-1-4939-0767-0.

[11] R. Brincker, C.E. Ventura, Introduction to Operational Modal Analysis, Wiley,

London, 2015. doi:10.1002/9781118535141.

[12] R. Brincker, L. Zhang, P. Andersen, Modal identification of output-only systems using frequency domain decomposition, Smart Mater. Struct. 10 (2001) 441–445. doi:10.1088/0964-1726/10/3/303.

[13] J.S. Bendat, A.G. Piersol, Engineering applications of correlation and spectral analysis., New York: Wiley, 1980., 1980.

[14] I. James, George H., T.G. Carne, J.P. Lauffer, The natural excitation technique (NExT) for modal parameter extraction from operating structures, Int. J. Anal. Exp. Modal Anal. 10 (1995) 260–277.

[15] B. Peeters, G. De Roeck, Reference-Based Stochastic Subspace Identification for Output-Only Modal Analysis, Mech. Syst. Signal Process. 13 (1999) 855–878. doi:10.1006/mssp.1999.1249.

[16] P. Van Overschee, B.L. De Moor, Subspace identification for linear systems: Theory—Implementation—Applications, Springer Science & Business Media, 2012.

[17] C. Devriendt, P. Guillaume, The use of transmissibility measurements in output-only modal analysis, Mech. Syst. Signal Process. 21 (2007) 2689–2696. doi:10.1016/j.ymssp.2007.02.008.

[18] W.-J. Yan, W.-X. Ren, An Enhanced Power Spectral Density Transmissibility (EPSDT) approach for operational modal analysis: Theoretical and experimental investigation, Eng. Struct. 102 (2015) 108–119. doi:10.1016/j.engstruct.2015.08.009.

[19] E.T. Jaynes, Probability Theory: The Logic of Science, Cambridge university press, 2003.

[20] J.L. Beck, Bayesian system identification based on probability logic, Struct. Control Heal. Monit. 17 (2010) 825–847. doi:10.1002/stc.424.

[21] K.V. Yuen, L.S. Katafygiotis, Bayesian time-domain approach for modal updating using ambient data, Probabilistic Eng. Mech. 16 (2001).

[22] L.S. Katafygiotis, K.V. Yuen, Bayesian spectral density approach for modal updating using ambient data, Earthq. Eng. Struct. Dyn. 30 (2001) 1103–1123. doi:10.1002/eqe.53.

[23] W.-J. Yan, L.S. Katafygiotis, A two-stage fast Bayesian spectral density approach for ambient modal analysis. Part I: Posterior most probable value and uncertainty, Mech. Syst. Signal Process. (2014) 1–17. doi:10.1016/j.ymssp.2014.07.027.

[24] W.-J. Yan, L.S. Katafygiotis, A two-stage fast Bayesian spectral density approach for ambient modal analysis. Part II: Mode shape assembly and case studies, Mech. Syst.

Signal Process. (2014) 1–16. doi:10.1016/j.ymssp.2014.08.016.

[25] K.V. Yuen, L.S. Katafygiotis, Bayesian fast Fourier transform approach for modal updating using ambient data, Adv. Struct. Eng. 6 (2003) 81–95.

[26] S.-K. Au, F.-L. Zhang, Y.-C. Ni, Bayesian operational modal analysis: Theory, computation, practice, Comput. Struct. 126 (2013) 3–14. doi:10.1016/j.compstruc.2012.12.015.

[27] J.L. Beck, L.S. Katafygiotis, Updating Models and Their Uncertainties. I: Bayesian Statistical Framework, J. Eng. Mech. 124 (1998) 455–461. doi:10.1061/(ASCE)0733-9399(1998)124:4(455).

[28] J.R. Vig, Introduction to Quartz Frequency Standards, (1992).

[29] D.L. Mills, Internet time synchronization: the network time protocol, IEEE Trans. Commun. 39 (1991) 1482–1493. doi:10.1109/26.103043.

[30] E. Kaplan, C. Hegarty, Understanding GPS: principles and applications, Artech house, 2005.

[31] J. Elson, L. Girod, D. Estrin, Fine-grained network time synchronization using reference broadcasts, ACM SIGOPS Oper. Syst. Rev. 36 (2002) 147–163. doi:10.1145/844128.844143.

[32] Y. Wang, J.P. Lynch, K.H. Law, Wireless structural sensors using reliable communication protocols for data acquisition and interrogation, Proc. 23rd Int. Modal Anal. Conf. (IMAC XXIII). (2005).

[33] S. Ganeriwal, R. Kumar, M.B. Srivastava, Timing-Sync Protocol for Sensor Networks, in: Proc. 1st ACM Int. Conf. Embed. Networked Sens. Syst., 2003: pp. 138–149. doi:10.1145/958507.958508.

[34] D.R. Brillinger, Time series: data analysis and theory, Siam, 1981.

[35] S. Bjorklund, L. Ljung, A review of time-delay estimation techniques, in: 42nd IEEE Int. Conf. Decis. Control (IEEE Cat. No.03CH37475), 2003: pp. 2502–2507. doi:10.1109/CDC.2003.1272997.

[36] G. Jacovitti, G. Scarano, Discrete time techniques for time delay estimation, IEEE Trans. Signal Process. 41 (1993) 525–533. doi:10.1109/78.193195.

[37] A. Demir, A. Mehrotra, J. Roychowdhury, Phase noise in oscillators: A unifying theory and numerical methods for characterization, Circuits Syst. I Fundam. Theory Appl. IEEE Trans. 47 (2000) 655–674. doi:10.1109/81.847872.

[38] M. Löhning, G. Fettweis, The effects of aperture jitter and clock jitter in wideband ADCs, Comput. Stand. Interfaces. 29 (2007) 11–18. doi:10.1016/j.csi.2005.12.005.

[39] Y.-C. Zhu, S.-K. Au, Spectral characteristics of asynchronous data in operational modal analysis, Struct. Control Heal. Monit. (2017) e1981. doi:10.1002/stc.1981.

[40] S.-K. Au, Fast Bayesian FFT method for ambient modal identification with separated modes, J. Eng. Mech. (2011). doi:10.1061/(ASCE)EM.1943-7889.0000213.